The composition and dynamic nature of the N-linked glycoprofile of bovine milk serum and its individual proteins

Valk-Weeber, Rivca

DOI:
10.33612/diss.134363958

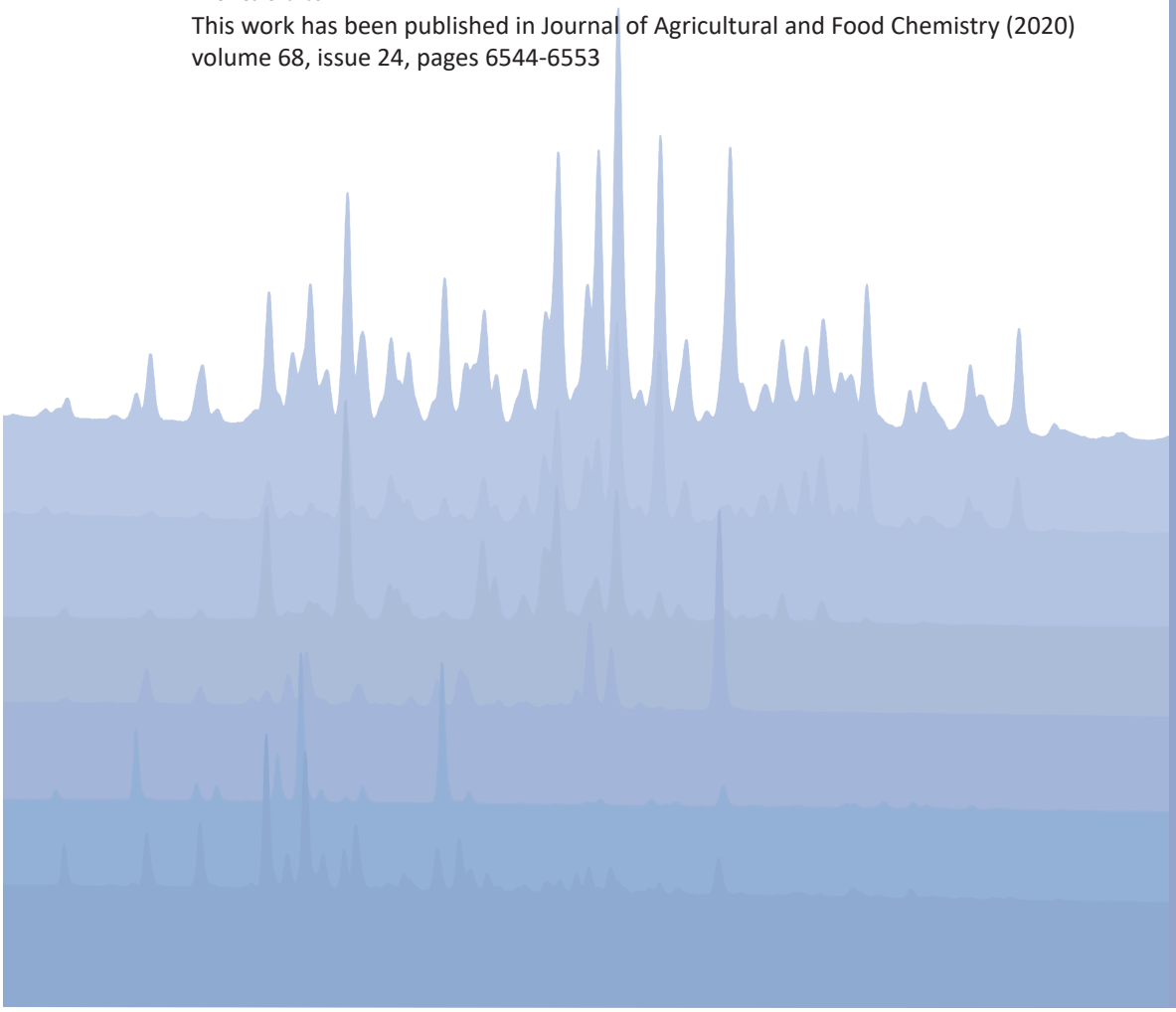
Chapter 6

In depth analysis of the contribution of specific glycoproteins to the overall bovine whey *N*-linked glycoprofile

Rivca L. Valk-Weeber[1], Cecile Deelman-Driessen[1,3], Lubbert Dijkhuizen[1,2], Talitha Eshuis-de Ruiter[4], and Sander S. van Leeuwen[1,5]

[1]Microbial Physiology Research Group, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Groningen, The Netherlands
[2]Current address: CarbExplore Research BV, Zernikepark 12, 9747 AN Groningen, The Netherlands
[3]Current address: Membrane Enzymology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB) University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
[4]Life Sciences, FrieslandCampina, Stationsplein 4, 3818 LE Amersfoort, The Netherlands
[5]Current address: Laboratory Medicine, University Medical Center Groningen (UMCG), Hanzeplein 1, 9713 GZ, Groningen, The Netherlands

This work has been published in Journal of Agricultural and Food Chemistry (2020) volume 68, issue 24, pages 6544-6553



Abstract

The *N*-linked glycoprofile of bovine whey is the combined result of individual protein glycoprofiles. In this work, we provide in-depth structural information on the glycan structures of known whey glycoproteins, namely lactoferrin, lactoperoxidase, α-lactalbumin, immunoglobulin-G (IgG) and glycosylation dependent cellular adhesion molecule 1 (GlyCAM-1, PP3). The majority (~95%) of *N*-glycans present in the overall whey glycoprofile were attributed to three proteins; Lactoferrin, IgG and GlyCAM-1. We identified specific signature glycans for these main proteins; Lactoferrin contributes oligomannose-type glycans, while IgG carries fucosylated di-antennary glycans with Gal-β(1,4)GlcNAc (LacNAc) motifs. GlyCAM-1 is the sole whey glycoprotein carrying tri- and tetra-antennary structures, with a high degree of fucosylation and sialylation. Signature glycans can be used to recognize individual proteins in the overall whey glycoprofile, as well as for protein concentration estimations. Application of the whey glycoprofile analysis to colostrum samples revealed dynamic protein concentration changes for IgG, lactoferrin and GlyCAM-1 over time.

Introduction

Milk is classically considered to be composed of three fractions: (butter)fat, casein and serum. The serum fraction, frequently called whey, contains the proteins that remain after removal of the caseins (Haug *et al.*, 2007).

Main high-abundance proteins of the whey fractions are α-lactalbumin, β-lactoglobulin, serum albumin, immunoglobulin G (IgG), Glycosylation dependent cellular adhesion molecule 1 (GlyCAM-1; also known as proteose peptone 3, PP3, lactophorin) and lactoferrin. Proteins present in medium abundance include the immunoglobulins IgA, IgM, lactoperoxidase and osteopontin (O'Riordan *et al.*, 2014b). Minor abundance proteins include lysozyme and folate binding protein, but also many others. In total over 900 minor abundance proteins have been identified, most of which have not been extensively studied (Tacoma *et al.*, 2015).

While some proteins are critical for milk stability (β-lactoglobulin) or for solubilizing calcium phosphate (caseins) (Farrell *et al.*, 2004), others are known to have specific biological functions. These bioactive proteins are often glycosylated, e.g. lactoferrin, immunoglobulins, lysozyme and lactoperoxidase. Decoration of proteins with carbohydrate moieties occurs either *N*- or *O*-linked, based on the location of the glycans. Mucin-type *O*-linked glycans, initiating with *N*-acetyl-galactosamine (GalNAc) bound to a serine or threonine residue, differ greatly from *N*-linked glycans, initiating with a tri-mannosyl-chitobiose core bound to an asparagine residue (Moremen *et al.*, 2012).

Some proteins carry exclusively *N*-linked (e.g. lactoferrin) or *O*-linked structures (e.g. osteopontin), while others may carry both (e.g. immunoglobulins, GlyCAM-1). Analysis of *O*- and *N*-linked structures requires different approaches. While most *N*-linked glycans can be released by peptide:*N*-Glycosidase F (PNGase F), no such universal enzyme is available for *O*-linked structures; their release typically involves chemical treatment, e.g. alkali β-elimination. Here we focus on isolation and structural analysis of *N*-linked glycans.

A number of studies have focused on individual bovine whey glycoproteins (Inagaki *et al.*, 2010a; van Leeuwen *et al.*, 2012a). These glycoproteins each have their own glycosylation fingerprints (arising from the glycan structures present). All glycoproteins in a milk sample contribute to its overall whey glycoprofile. Glycan structures of similar size and monosaccharide composition tend to co-elute in chromatographic analysis. This complexity makes identification of individual glycan structures a challenge. The types of glycans present on most bioactive whey glycoproteins have been annotated (O'Riordan *et al.*, 2014b). It is unknown, however, what the contribution is of individual glycoproteins to

the overall whey glycoprofile. Changes in whey protein glycans over the course of lactation have been reported, but these studies focused mostly on IgG and lactoferrin, leaving GlyCAM-1 unstudied (Takimori *et al.*, 2011).

Whey protein powders, containing IgG, lactoferrin and GlyCAM-1, are processed into different food products, including infant formulas (Mettler, 1980). Lactoferrin is known to have antimicrobial and immunostimulatory functions. The latter function is mediated by Toll-like receptors, and depends on the composition of the lactoferrin glycoprofile (Figueroa-Lozano *et al.*, 2018). Similarly, core-fucosylation as present on the glycans of IgG is crucial for receptor interaction (Takahashi *et al.*, 2009). Unique functions for GlyCAM-1 and its glycans remain to be identified, although evidence exists for antimicrobial and mucin-like lubricating properties of this protein (Campagna *et al.*, 2004; Dowbenko *et al.*, 1993). Efficient methods for the unraveling of the overall glycoprofile of whey are crucial for predicting the functional properties of whey, and the products they are processed into.

Here, we used UPLC-FLD to identify unique signature *N*-glycans of the whey proteins lactoferrin, lactoperoxidase, α-lactalbumin, IgG and GlyCAM-1. In addition, an overview of the *N*-glycan contribution of each protein to the overall whey glycoprofile is provided. We applied the overall whey glycoprofile analysis method towards milk and colostrum samples. Information on the *N*-glycans of lactoferrin and IgG and their protein concentration in colostrum already was available (Takimori *et al.*, 2011; Valk-Weeber *et al.*, 2020a), but this information was lacking for GlyCAM-1. Here we show that the concentrations of lactoferrin, IgG and GlyCAM-1 in whey can be followed over time by analysis of their unique glycan structures in the whey glycoprofile.

Materials and methods

Materials

Bovine lactoferrin, lactoperoxidase and α-lactalbumin samples were provided by FrieslandCampina Domo (Amersfoort, the Netherlands). Bovine gamma globulin fraction 2 (purity > 98%) was from Serva (Heidelberg, Germany). PNGase F (*Flavobacterium meningosepticum*) was from New England Biolabs (Ipswich, UK). Jack bean α-mannosidase (75 U/mL in 3.0 M $(NH_4)_2SO_4$, 0.1 mM zinc acetate, pH 7.5) was purchased from Sigma-Aldrich Chemie N.V. (Zwijndrecht, the Netherlands). Green coffee bean α-galactosidase (25 U/mL 100 mM sodium phosphate pH 6.5, containing 0.25 mg/ml bovine serum albumin), bovine testis β-galactosidase (5 U/mL in 20 mM sodium citrate phosphate, 150 mM NaCl, pH 4.0), *Streptococcus pneumoniae* sialidase, (4 U/mL 20 mM Tris-HCl, 25 mM NaCl, pH 7.5) and *Arthrobacter ureafaciens* α-sialidase (5 U/mL in 20 mM Tris-HCl pH 7.5, containing 25 mM NaCl) were from Prozyme (Ballerup, Denmark).

Streptococcus pneumoniae β-*N*-acetylhexosaminidase (40 U/mL in 20 mM Tris-HCl, 50 mM NaCl pH 7.5), jack bean β-*N*-acetylhexosaminidase (50 U/mL in 20 mM sodium citrate phosphate pH 6.0), bovine kidney α-fucosidase (2 U/mL in 20 mM sodium citrate phosphate, 0.25 mg/ml BSA pH 6.0 were from Prozyme (Ballerup, Denmark). Pooled tank milk of Holstein-Friesian cows was obtained from FrieslandCampina Domo. Colostrum and milk samples were collected from 8 cows from a local organic farm (Rietveldhoeve farm, Aduard, Groningen, the Netherlands). Colostrum was collected directly after calving, and then at approximately 12, 24, 36, 48, 60 h *post-partum* (Table S1). Milk samples were collected at 1, 2 and 3 months.

Whey preparation and protein isolation

Milk samples were thawed in a water bath at 37 °C and homogenized. An aliquot of 1 mL was defatted by centrifuging at 4,000 x *g* for 10 min. An amount of 400 µL defatted milk was transferred into a new tube. Of colostrum samples, an amount of 50 µL was transferred and mixed with 350 µL of MilliQ water. For GlyCAM-1 analysis, defatted milk was heated to 95°C for 30 min prior to subsequent processing. Caseins were removed by addition of 400 µL 125 mM of ammonium acetate at pH 4.6 (ratio of 1:1). The samples were vortexed and left at room temperature for 5 min before centrifuging at 11,000 x *g* for 5 min to precipitate the caseins. An aliquot of 100 µL of the supernatant (acid whey) was transferred into a new tube, 400 µL 100 mM ammonium acetate in methanol (MeOH+NH_4Ac) was added and mixed by vortexing. Whey protein precipitation was facilitated by centrifugation for 5 min at 11,000 x *g*. The solvent (containing lactose) was carefully pipetted from the protein pellets. The protein pellets were re-dissolved in 75 µL of 2% SDS and 2% β-mercaptoethanol in 80 mM phosphate buffer at pH 7.5. After addition of the solvent, the samples were incubated at 37 °C for 10 min, after which they were vortexed vigorously and further incubated for an additional 10 min, followed by a final vortex mixing. The proteins were denatured for 15 min at 85 °C and cooled to room temperature. An aliquot of 25 µL of 10% NP-40 (NP-40 substitute, Sigma) was added to each sample and mixed by vortexing. Finally, 2 µL of diluted PNGase F (100 units/experiment) was added to the samples and mixed. Glycans were released overnight at 37 °C.

Labeling and cleanup

Isolated glycans were labeled with anthranilic acid (2-AA, Sigma) or 2-aminobenzamide (2-AB). The 2-AA label was chosen for applications with fluorescent detection due to the higher sensitivity in these applications. 2-AB was chosen for applications that required mass spectrometry analysis. Direct in solution labeling of whey digests was performed as follows. Whey protein digests of a total volume of 102 µL were mixed 1:1 with labeling solution (0.7 M 2-AA or 2-AB and 2 M of 2-picoline borane or sodium cyanoborohydride in

dimethylsulfoxide (DMSO, Sigma): glacial acetic acid (7:3, v/v)). Incubations were performed for 2 h at 65 °C (Bigge *et al.*, 1995). Labeling reagents were removed by 96-well microcrystalline cellulose SPE as described (Ruhaak *et al.*, 2008). Samples were diluted with 612 µL acetonitrile (final concentration 75% v/v) prior to application to the cellulose SPE.

Anion-exchange SPE fractionation

For fractionation of the glycans into sialylated and neutral fractions, 4 aliquots of 100 µL of acid whey were processed and labeled with 2-AB as described above. The labeled aliquots were pooled and fractionated by anion exchange solid phase extraction (IRIS MAX, 1 mL, Screening Devices). The cartridge was conditioned with 1 mL of acetonitrile and 1 mL of MilliQ water. The labeled whey was passed over the column and collected (neutral glycan fraction). After washing with 1 mL of MilliQ, the acidic glycans were eluted with 1 mL of 25% acetonitrile+0.1% TFA. The fractions were lyophilized and redissolved in 400 µL of MilliQ before HPLC analysis.

HPLC analysis

Fluorescently labeled glycans were separated on an Acquity UPLC Glycan BEH Amide column (2.1 mm x 100 mm, 1.7 µm, Waters Chromatography BV, Etten-Leur, the Netherlands), using an UltiMate 3000 SD HPLC system (Thermo Fisher Scientific, Waltham, MA) equipped with a Jasco FP-920 fluorescence detector (λex 330 nm, λem 420 nm, Jasco Inc, Easton, MD). An injection volume of 3 µL was used. Ternary gradients were run using MilliQ water, acetonitrile and a buffer solution consisting of 250 mM formic acid in MilliQ water, adjusted to pH 3.0 using ammonia. A constant 20% of the buffer was maintained throughout the run. Elution was performed by a slow sloping gradient of 22% to 40% MilliQ water (total concentration, including buffer) from 0 to 67.5 min. The remaining percentage of the solvent composition comprised of acetonitrile. After completion of the gradient, final gradient conditions were maintained for 9 min and the column reconditioned back to initial conditions for 13 min.

Exoglycosidase assays

Sequential digestions with glycosidases (supplemental material) were performed in 50 mM sodium acetate buffer at pH 5.5 overnight. After each digestion step, the enzymes were removed by 10 kDa cut-off centrifugal filters (Millipore, Tullagreen, Cork, IRL). The 2-AA labeled dextran calibration ladder was from Waters Chromatography BV (Etten-Leur, the Netherlands).

Lactoferrin concentration determination

The concentration of lactoferrin in the colostrum samples was quantified by a bovine lactoferrin ELISA quantitation set (E10-126, Bethyl Laboratories, Montgomery, TX, USA), as described previously (Valk-Weeber *et al.*, 2020).

Mass Spectrometry Analysis

Mass spectrometry analysis was performed using identical slope and solvent composition as used for the HPLC-fluorescent detection. Settings used for the mass spectrometry analysis were as described earlier (Valk-Weeber *et al.*, 2020). Glycans were identified by their (derivatized) monoisotopic molecular mass, using the GlycoMod tool (Cooper *et al.*, 2001) (https://web.expasy.org/glycomod/) and a 0.2 Dalton mass tolerance.

Results and discussion

Overall whey glycoprofile

The overall *N*-glycan profile of bovine acid whey showed a complex pattern of peaks (Fig. 1). There are multiple glycosylated proteins present in acid whey that contribute *N*-glycans to the overall chromatogram. Glycans with similar degree of polymerization and monosaccharide composition tend to elute at the same time. Due to the high number of structures present in the chromatogram, multiple structures can overlap and form combined peaks (peak clusters). Structures, including their isomers were identified by LC-MS. It should be noted that for optimal fluorescent detection, glycans were labeled with 2-AA, while for LC-MS analysis glycans were labeled with 2-AB for improved positive ion mode sensitivity. Identical chromatographic conditions were used for both detection methods. While glycans labeled with 2-AB have a higher retention in the chromatography setup used, the chromatographic patterns are the same (Valk-Weeber *et al.*, 2020). Using the structures identified in the 2-AB labeled glycoprofile, the structures in the 2-AA labeled whey glycoprofile were appointed (Fig. 1). Structures were further confirmed by sequential exoglycosidase treatment (Figures S1 and S2). In our study, we were able to identify at least 69 individual glycan structures, not including isomers (see overview in Table S2).

Sialylated and neutral glycans were separated by anion-exchange SPE and profiled (Fig. 2). The neutral glycans (Fig. 2, black line, 15-33 min) dominated the first half of the chromatogram, while the sialylated glycans eluted in the second half (Fig. 2, red line, 33-55 min). Shorter sialylated structures overlap with the larger neutral structures between 33- and 47-min retention.

Previous publications have reported the absence of α(2,3)-linked sialic acid on bovine whey glycans (Takimori *et al.*, 2011; van Leeuwen *et al.*, 2012b), while others did not specify the linkage type (Sriwilaijaroen *et al.*, 2012). Recently, we reported the presence of α(2,3) sialic acid on bovine lactoferrin isolated from colostrum, but not on the mature milk-derived protein (Valk-Weeber *et al.*, 2020). Exoglycosidase treatment of the (mature) whey glycoprofile, with sialidase from *Streptococcus pneumoniae*, with strong preference for α(2,3)-linked sialic acid,

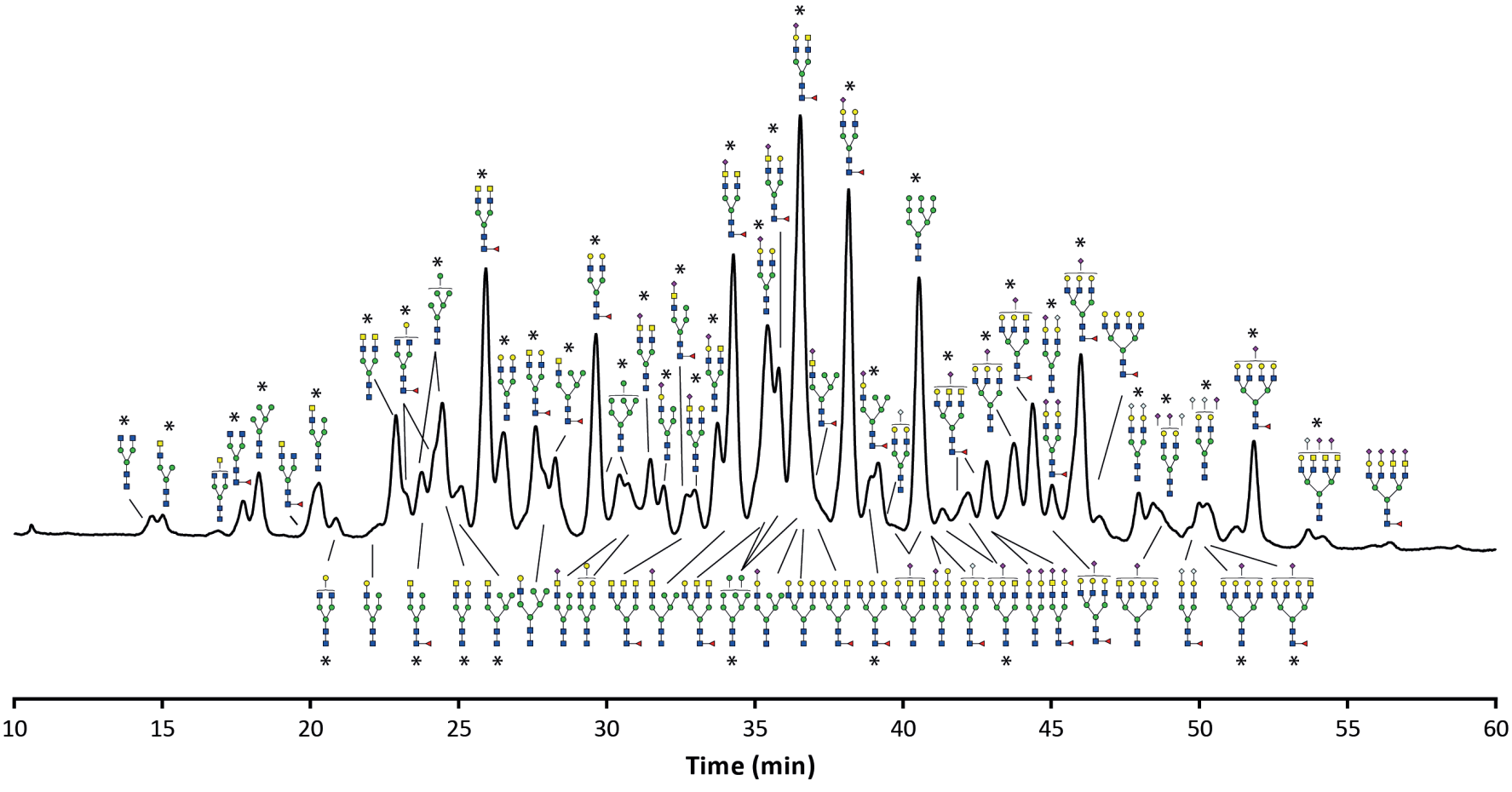

10
15
20
25
30
35
40
45
50
55
60
Time (min)

Figure 1 (left). Overall acid whey HPLC glycoprofile obtained for pooled milk from Holstein-Friesian cows. Glycan structures detected and identified by mass spectrometry were added to the spectrum; for a full overview of all structures, see Table S2. Structures with main contributions to the peak intensity are marked with *.

confirmed the presence of α(2,3) linkages in trace amounts (Fig. S3). This sialic acid was present on multiply sialylated di-, tri- and tetra-antennary structures, not only on lactoferrin-derived structures. This indicates that the presence of α(2,3)-linked sialic acid is a general feature of whey glycoproteins, and not only of lactoferrin in the colostrum phase.

The bovine milk glycoprofile has been investigated previously, either with glycans isolated from commercial whey powders (van Leeuwen *et al.*, 2012b), colostrum whey (Karav *et al.*, 2015b) or mature milk (Holstein and Jersey cows) (Nwosu *et al.*, 2012)

In the analysis of bovine whey glycosylation, the *N*-acetylgalactosamine-*N*-acetylglucosamine (LacdiNAc) motif is important to consider, especially when mass spectrometry is used for structural identification. Since *N*-acetyl-glucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc) are isomers and cannot be distinguished by mass spectrometry, MS-based annotation is often difficult and leads to ambiguous results. However, considering the high amount of LacdiNAc motifs reported on bovine milk proteins in earlier studies (Sriwilaijaroen *et al.*, 2012; van Leeuwen *et al.*, 2012b), the presence of LacdiNAc motifs in high abundance was expected. Karav *et al.* and Nwosu *et al.*, made no distinction between GlcNAc and GalNAc and instead, the shared identifier HexNAc was used. These papers do not report LacdiNAc structures, but refer to non-galactosylated tri- and tetra-antenna structures instead (Karav *et al.*, 2015b; Nwosu *et al.*, 2012). Moreover, whereas some studies report complex-type structures up to tetra-antennary with up to three sialic acid residues (Takimori *et al.*, 2011; van Leeuwen *et al.*, 2012b), other studies report a more limited glycoprofile, mainly oligomannose and di-antennary complex-type structures. Sriwilaijaroen *et al.* used PNGase A (instead of PNGase F), which has affinity for oligomannose-, hybrid- and short complex-type (up to di-antennary) glycans, possibly explaining the more limited glycoprofile obtained (Sriwilaijaroen *et al.*, 2012). Karav *et al.* used the bifidobacterial enzyme EndoBI-1 instead of PNGase F (Karav *et al.*, 2015b). This enzyme cleaves between the two GlcNAc residues of the chitobiose core, thereby information on core-fucosylation is lost. Previous reports have shown that a significant number of structures carry core-fucosylation (Nwosu *et al.*, 2012; Takimori *et al.*, 2011; van Leeuwen *et al.*, 2012b). Loss of this highly relevant information thus is a significant disadvantage of the use of the EndoBI-1 enzyme for this type of analysis. In our study, multiply sialylated di- and tetra-antennary structures were detected, but not multiply sialylated tri-antennary structures. Van Leeuwen *et al.* (2012b) detected trace amounts of multiply sialylated tri-antennary structures, but only in concentrated

fraction; these structures may have remained below the limit of detection in our study. Conversely, we detected a number of doubly sialylated tetra-antennary structures, not reported by van Leeuwen *et al.* Therefore, the data found in our study and in earlier work complement each other. Overall, the level of complexity observed in our study is comparable with that in Takimori *et al.* and van Leeuwen *et al.* (Takimori *et al.*, 2011; van Leeuwen *et al.*, 2012b).

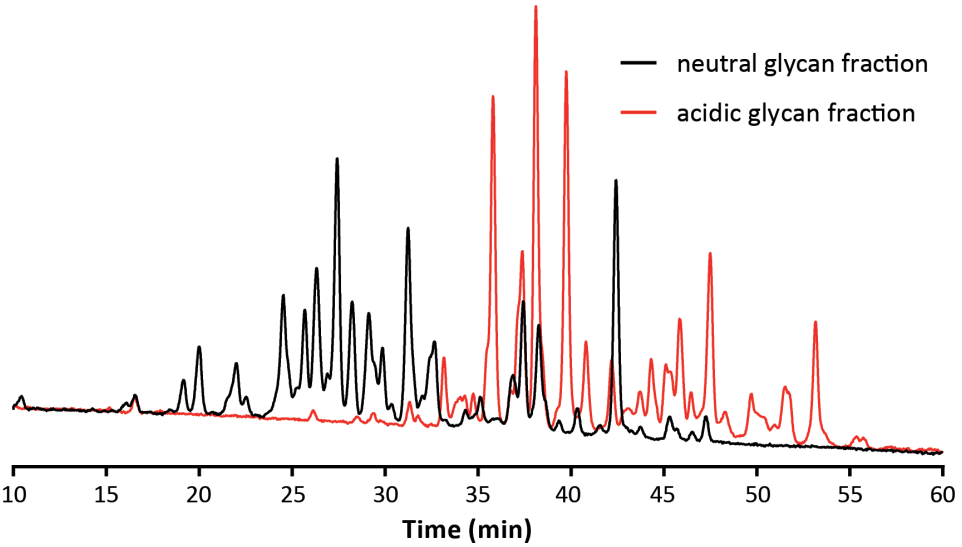


Figure 2. Overlay of the neutral (black line) and acidic (sialylated, red line) glycan fractions of the overall whey HPLC glycoprofile, obtained from pooled milk of Holstein-Friesian cows. The 2-AB labeled glycans were divided into neutral and acidic fractions by anion exchange chromatography fractionation.

Individual whey glycoproteins

Glycoproteins in bovine whey each have a signature fingerprint of glycans. The concentration of these proteins in bovine whey varies, highest concentrations were reported for IgG (0.3-0.6 mg/mL), GlyCAM-1 (0.3-0.5 mg/mL) and lactoferrin (0.1-0.3 mg/mL) (Farrell *et al.*, 2004; Larson & Rolleri, 1955). The other immunoglobulins, IgA and IgM are present at approximately 5 to 10 times lower concentrations than IgG (Korhonen *et al.*, 2000). While α-lactalbumin is present in higher protein concentrations (1.5 mg/mL), only ~10% of this protein is glycosylated (Slangen & Visser, 1999). Lactoperoxidase is typically present in concentrations around 0.03 mg/mL (Kussendrager & van Hooijdonk, 2000). Together, these proteins are most likely the main contributors to the overall whey *N*-linked glycoprofile.

GlyCAM-1 is currently not commercially available and therefore had to be isolated from milk samples. When milk is heated, only the heat-stable proteins remain in solution. Acid whey prepared of this heated milk is known as the proteose peptone (PP) fraction, with GlyCAM-1 initially labeled as PP3 (Larson *et al.*, 1955). PP3 reportedly is the main contributor to the *N*-linked glycoprofile (Sørensen

& Petersen, 1993b); the second most abundant protein is a casein proteolytic fragment (PP5) that is not *N*-glycosylated (Vreeman *et al.*, 1986). The 60 kDa protein osteopontin can also be recovered from the PP fraction. Osteopontin, like casein, is solely *O*-glycosylated and will therefore not interfere in our *N*-glycoprofile analysis (Sørensen *et al.*, 1995). SDS-PAGE analysis confirmed that other glycoproteins (IgG/lactoferrin/α-lactalbumin) were sufficiently removed by heating the whey (Fig. S4). Glycan fingerprints were analyzed for the selected main glycoproteins as well as for heated whey (GlyCAM-1) (Fig. 3).

The glycan fingerprint of GlyCAM-1 is dominated by sialylated complex-type glycan structures, existing in di-, tri-, and tetra-antennary configurations (Fig. 3). The observed high levels of sialylated and core-fucosylated glycans for GlyCAM-1 fit previous reports (Inagaki *et al.*, 2010a). A few structures were observed here that have not been previously described for GlyCAM-1. For example, we observed multiply sialylated structures, in some cases with a combination of Neu5Ac and Neu5Gc. These include a tetra-antennary structure with two Neu5Ac and one Neu5Gc moiety (Table S2, nr 70). In addition, tri-sialylated di-antennary structures were found (Table S2, nr 63-65), indicating the addition of a third sialic acid on a GlcNAc, instead of on the terminal Gal(NAc). It should be noted that sialylation on GlcNAc only occurrs via an α(2,6) linkage, and only when the Gal(NAc) is α(2,3) sialylated (Stanley & Cumming 2017).

The glycans on α-lactalbumin showed significant overlap with the glycans found on GlyCAM-1 (Fig. 3). The glycan fingerprint of α-lactalbumin was characterized by the presence of di-antennary glycans, of which the majority bears the GalNAc-β(1,4)-GlcNAc (LacdiNAc) motif. Both fucosylation and sialylation was abundantly present on the glycans from α-lactalbumin. A minor amount of tri-antennary structures was also identified. The major structures found in the glycosylation fingerprint of α-lactalbumin in this study corresponded with those in an earlier report (Slangen *et al.*, 1999). Minor additional peaks were also observed in the glycan fingerprint, indicating the presence of additional glycan structures. Based on their positions in the chromatogram, these other glycan structures are hypothesized to be variations of the identified di-antennary and tri-antennary structures.

The glycan fingerprint of lactoferrin was dominated by the oligomannose structures Man-5 to Man-9, with Man-8 and Man-9 being the most abundant of the set (Fig. 3), fitting previous results on mature milk derived lactoferrin (Valk-Weeber *et al.*, 2020; van Leeuwen *et al.*, 2012a). In addition, di-antennary structures were present, decorated with either galactose or *N*-acetylgalactosamine. Hybrid type structures were also found on lactoferrin. The complete profile of glycan structures of bovine lactoferrin fits to previous reports (Spik *et al.*, 1988; Valk-Weeber *et al.*, 2020; van Leeuwen *et al.*, 2012a).

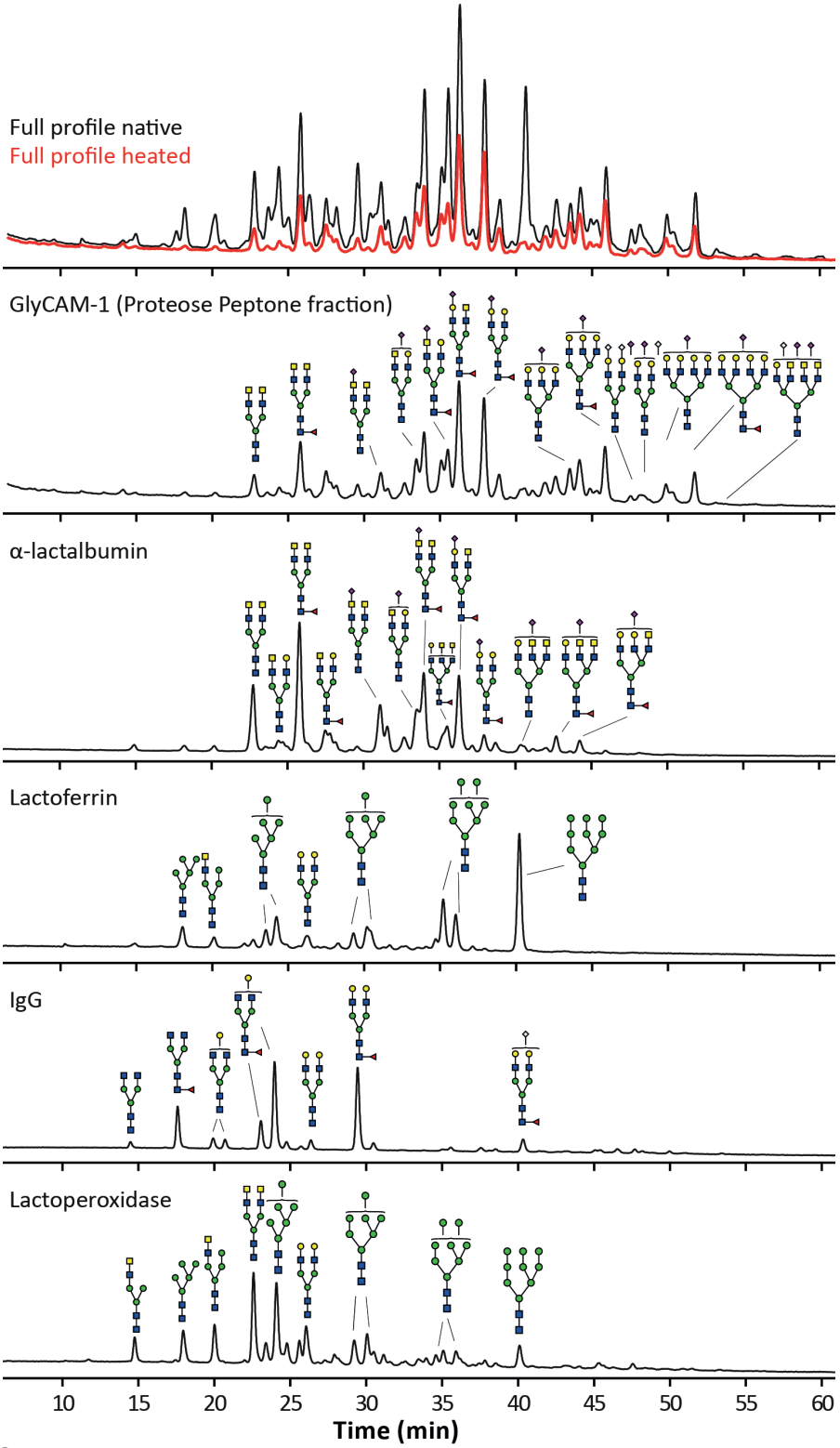

Full profile native
Full profile heated
GlyCAM-1 (Proteose Peptone fraction)
α-lactalbumin
Lactoferrin
IgG
Lactoperoxidase
10
15
20
25
30
35
40
45
50
55
60
Time (min)

Figure 3 (left). HPLC chromatograms of the glycan fingerprints of individual glycoproteins present in bovine acid whey. Glycan structures of the most abundant peaks were annotated.

In earlier work, Wolf *et al.* identified the glycan structures on lactoperoxidase (Wolf *et al.*, 2000). Our study visualized the same glycans (Fig. 3), and also confirmed the presence of hybrid structures, which were hypothesized by Wolf *et al.* In addition, the relative quantities of the individual glycans of lactoperoxidase could be calculated. Lactoperoxidase carries a mixture of oligomannose, di-antennary complex and hybrid structures. The relative abundance of the oligomannoses on lactoperoxidase differs from that observed on lactoferrin. On lactoperoxidase, the oligomannoses Man-5 to Man-7 were most abundant. Hybrid- and complex-type structures are decorated with galactose or GalNAc, with the doubly GalNAc decorated structure present in the highest amounts. A small quantity of sialylated structures was also detected on lactoperoxidase, decorated with Neu5Ac or Neu5Gc, or a combination of these sialic acids.

Bovine IgG contains di-antennary glycan structures, of which the majority was core-fucosylated (Fig. 3). The antennae were decorated with galactose (LacNAc), which is unique as all other bovine milk glycoproteins contain significant amounts of LacdiNAc epitopes. A number of structures was sialylated; Neu5Gc is the predominant sialic acid on this protein.

Signature glycans and contribution of individual proteins to the overall whey glycoprofile

The individual protein glycan fingerprints gave valuable information on the glycan heterogeneity of each protein. Next, we attempted to visualize the contribution of each protein to the overall whey glycoprofile. An overlay of the glycan fingerprints of the individual whey proteins was prepared, reflecting their reported concentrations in a bovine whey sample (Fig. 4). The concentrations were chosen according to the literature established concentration range for IgG, lactoferrin, α-lactalbumin and lactoperoxidase. GlyCAM-1 was analyzed directly from a heated defatted milk sample (pooled tank milk) and therefore represents a typical milk concentration. In the resulting overall glycoprofile, many glycan structures overlap. However, various individual glycans could be identified as signature structures for a single glycoprotein.

The majority of the glycan structures in the overall whey glycoprofile appeared to originate from GlyCAM-1 (Fig. 4, blue line). A large portion of the glycans from GlyCAM-1 was sialylated, and the majority of the sialylated structures in the whey glycoprofile are likely to originate from GlyCAM-1. The acidic (sialylated) glycan fraction of the glycan pool indeed showed striking similarities to the glycan fingerprint of GlyCAM-1 (compare Figures 2, 4). Tri- and tetra-antennary structures were not observed (or in very minor quantities) on the other glycoproteins. Therefore, tri- and particularly tetra-antennary structures (with

and without sialylation) are signature glycan structures for GlyCAM-1. While sialylated di-antennary structures were also observed on the other glycoproteins, the majority of these glycans originate from GlyCAM-1. LacdiNAc motifs are very common on α-lactalbumin and GlyCAM-1, however the contribution of α-lactalbumin to the overall glycoprofile is limited (see Figure 4 and below).

The second highest contributor to the overall bovine whey glycoprofile is lactoferrin. The oligomannose type glycans that were found in the overall glycoprofile are almost exclusively from lactoferrin (Fig. 4). While this glycan class was also found on lactoperoxidase (Fig. 3), the contribution of lactoperoxidase to the overall glycoprofile was minimal (see Figure 4 and below). Man-9 in particular eluted in a relatively isolated part of the overall glycoprofile (Fig. 4, starred structure), with little overlap or contribution from other co-eluting glycan structures. Therefore, the Man-9 peak was identified as signature glycan for lactoferrin in milk and whey (Fig. 4).

The third highest contributor to the overall glycoprofile is IgG. Core fucosylated di-antennary glycans with LacNAc motifs are signature glycans for IgG. The glycans from IgG eluted in the first half of the chromatogram (Fig. 4). The three most abundant glycan structures of IgG were readily identifiable in the overall whey glycoprofile, although there was some overlap with other co-eluting species (Fig. 4).

As described earlier, only ~10% of the total amount of α-lactalbumin is glycosylated, therefore, the contribution of this protein, although present at a relatively high concentration, to the overall glycoprofile remained very low (Fig. 4). The glycan structures found on α-lactalbumin were also present on GlyCAM-1 (Fig. 3). Therefore, no unique signature structures were identified for α-lactalbumin.

Although lactoperoxidase is clearly glycosylated (Fig. 3), it does not possess any unique identifier glycan structures. Especially lactoperoxidase and lactoferrin showed similarities in glycan structures, albeit in different relative quantities (Fig. 3). In view of the low concentration of lactoperoxidase in bovine whey, its contribution to the overall glycoprofile of whey is nihil (Fig. 4). No signature glycans were identified for lactoperoxidase.

Other immunoglobulins

While the protein concentrations of IgA and IgM are ~10 times lower than IgG protein, their percentages of carbohydrate are higher than for IgG. For IgG, the carbohydrate content was estimated at 2-4%, consisting entirely of *N*-glycans (Butler, 1969; Deisenhofer, 1981). For IgA the carbohydrate content was estimated at 7-10%, while IgM contains 10-12% carbohydrate per weight (Butler, 1969). Although present in much lower protein concentrations, IgA and IgM may still contribute significantly to the overall whey glycoprofile. To the best of our knowledge, no glycoprofiling for bovine IgA and IgM has been performed.

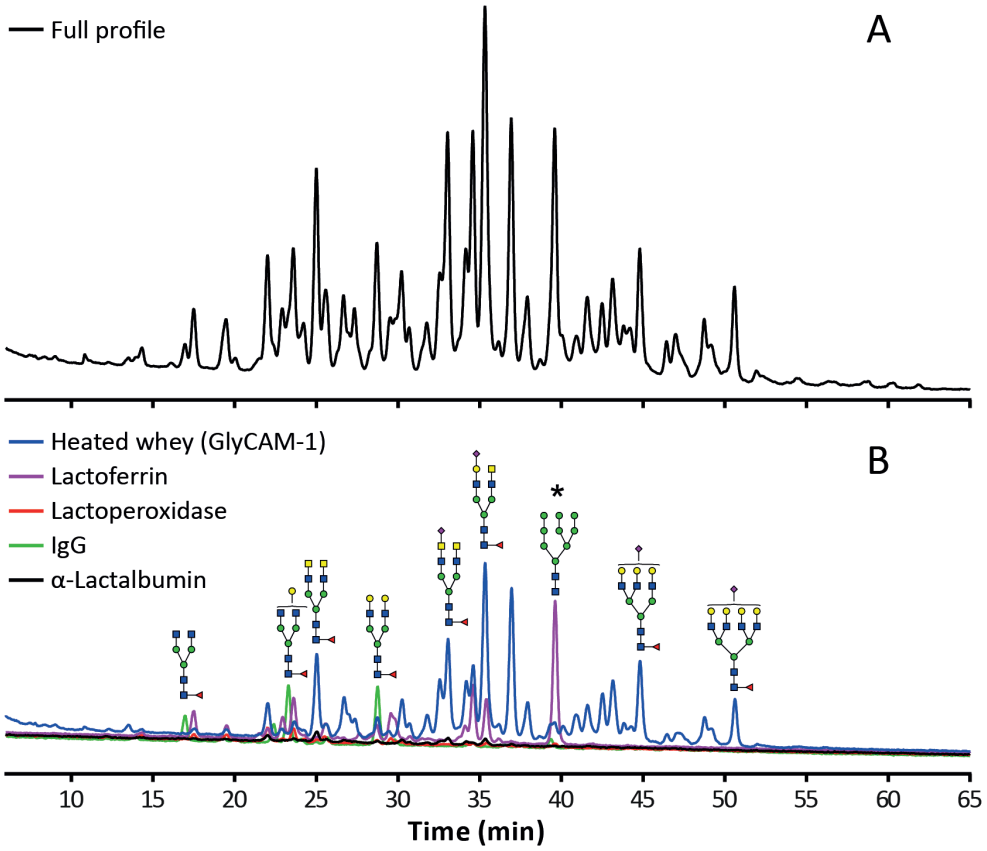


Figure 4. HPLC chromatograms of the glycan fingerprints of the individual whey glycoproteins shown at their reported physiological concentrations. The overall whey glycoprofile is shown at the top (A)(undiluted), and the individual glycan fingerprints at the bottom (B). The proteins were overlaid in the following concentration ratios: heated whey (undiluted)(GlyCAM-1), lactoferrin (0.20 mg/mL), lactoperoxidase (0.05 mg/mL), IgG (0.40 mg/mL) and α-lactalbumin (1.5 mg/mL). A selection of structures of the individual proteins was annotated, for the full annotation, refer to Figure 1. The main glycan of lactoferrin, Man-9, is shown as a starred structure.

Human IgA contains *O*-glycans in addition to *N*-glycans, which contribute to the overall glycan weight, but not to the *N*-glycoprofile. The *N*-linked glycans of human IgA are of the di-antennary type (Baenziger & Kornfeld, 1974). Human IgM contains 5 *N*-linked glycosites on each heavy chain, occupied with di-antennary (77% of total) and oligomannose (23% of total) type glycans (Arnold *et al.*, 2005). Assuming that the glycosylation of bovine IgA and IgM is similar to that of their human variants, a low to moderate contribution to the di-antennary pool can be expected. But no unique glycan signature structures are evident. Further analysis of the glycoprofiles of bovine IgA and IgM is needed to draw solid conclusions.

Whey glycoprofiles of colostrum

In previous work, we have shown that the glycoprofile of lactoferrin undergoes significant alterations during the short colostrum period (Valk-Weeber *et al.*, 2020). Here, the whey glycoprofiles of colostrum and mature milk samples from eight different cows (Table S1) were analyzed and compared (Fig. 5). The protein content (both caseins and whey proteins) of colostrum is significantly higher than of mature milk (McGrath *et al.*, 2016). To allow for efficient casein coagulation and whey protein analysis, the colostrum was diluted 8 times prior to casein precipitation. To compare late colostrum with mature milk, the last colostrum sample was also analyzed without additional dilution. Two cows were selected for a comparative analysis, cow 1 and cow 3. Cow 1 had low concentrations (< 0.1 mg/mL) of lactoferrin during the colostrum phase, as determined by ELISA analysis (Table S1). In contrast, lactoferrin concentrations were very high (> 20 mg/mL) in the day 1 sample of cow 3.

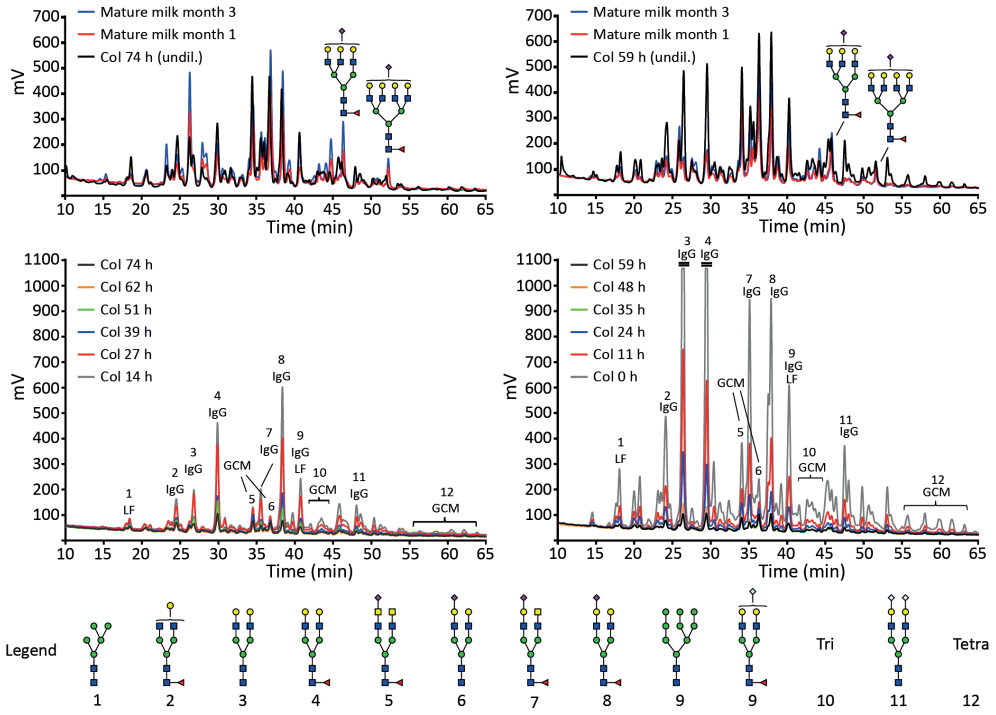


Figure 5. HPLC chromatograms of whey glycoprofiles of 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top). Glycoprofiles displayed are from (left) cow 1 and (right) cow 3 (Table S1). A selection of structures from lactoferrin (LF), immunoglobulin G (IgG), and GlyCAM-1 (GCM) was annotated. Sections with multiple co-eluting tri-antennary (Tri) and tetra-antennary (Tetra) glycans are bracketed. Additional glycoprofiles from cows 2, 4, 6, 7 and 8 are provided in the supplemental material (Figures S5-10).

In colostrum, large increases in the concentrations of lactoferrin and IgG were expected (McGrath *et al.*, 2016; Valk-Weeber *et al.*, 2020a). The difference in protein concentration of the major proteins (lactoferrin, IgG and GlyCAM-1) between these two cows also was reflected in their glycoprofiles: the colostrum glycoprofile for cow 3 was much more intense than the one obtained for cow 1 (Fig. 5). Based on the signature structures defined above, the galactosylated (LacNAc) di-antennary glycans mostly belonged to IgG, while oligomannoses mostly originated from lactoferrin. Sialylated di-antennary glycans with LacdiNAc motifs, tri- and tetra-antennary glycans are signature structures for GlyCAM-1. A selection of the signature structures was annotated in Figure 5. Using these signature glycan structures, it is notable that in cow 3 the structures of lactoferrin were present in higher concentrations than in cow 1, which is in agreement with the higher concentration of lactoferrin found by ELISA (Table S1). Moreover, the levels of IgG related glycans in cow 3 were significantly higher in the early colostrum phase than in cow 1. Based on the intensity of their signature glycan structures, both cows showed very rapid decreases in lactoferrin and IgG concentrations between the colostrum at day 1 and day 3. From the intensity of the signature structures of IgG, lactoferrin and GlyCAM-1, as well as from the relative proportion of the signature structures of these proteins, their relative protein concentrations were assessed. In early colostrum it appeared that the whey protein balance heavily shifted towards lactoferrin and IgG. Considering GlyCAM-1, a higher concentration of this protein was observed in colostrum, with a rapid decrease in concentration over the colostrum period. However, the concentration of GlyCAM-1 did not increase as extensively as that of IgG and lactoferrin.

The glycoprofiles obtained in early colostrum were different from those obtained from late colostrum and mature milk. Altered glycoprofiles, most notable by an increased degree of sialylation and fucosylation in early colostrum, have been reported for both IgG and lactoferrin (O'Riordan *et al.*, 2014a; Takimori *et al.*, 2011; Valk-Weeber *et al.*, 2020a). This was clearly reflected by the presence of high levels of sialylated IgG structures (Fig. 5, peaks 7, 8, 9, 11), which were absent, or severely decreased on mature IgG (Fig. 3). Care has to be taken to identify the tri-antennary glycan structures of GlyCAM-1: some of the upregulated sialylated di-antennary structures of IgG co-elute with the tri-antennary structures of GlyCAM-1 in colostrum (Fig. 5, area 10). The number of multiply sialylated tetra-antennary structures of GlyCAM-1 appeared higher in early colostrum, indicating that an increased sialylation was also occurring on GlyCAM-1 (Fig. 5, area 12).

Maturation of GlyCAM-1 appeared relatively slow in 50% of the analyzed cows. In cow 3, an increase in GlyCAM-1 signature structures was observed from month 1 to 3. Similar changes were observed for cows 2, 6 and 7 (Figures S5, 8, 9). This may indicate that the concentration of GlyCAM-1 increased during the first three months

of lactation, or that the glycosylation pattern had not stabilized completely, leading to an observed increase of less complex glycan structures later in the lactation cycle.

To our knowledge, GlyCAM-1 has not yet been quantified in bovine colostrum. In koala, a significantly higher GlyCAM-1 concentration was observed in colostrum versus mature milk (Morris *et al.*, 2016). In camel milk, GlyCAM-1 (PP3) was only detected in colostrum after 48 h post-partum (El-Hatmi *et al.*, 2007). Our results suggest that the GlyCAM-1 concentration in cows is increased in colostrum, followed by an immediate decrease and finally a slow stabilization over the first months of lactation. However, further research towards colostrum GlyCAM-1, both in concentration and glycosylation, is needed.

The contribution of GlyCAM-1

The reported concentration of GlyCAM-1 in bovine milk (0.3-0.4 mg/mL) (Larson *et al.*, 1955), was in the same range as that of the biologically relevant proteins IgG (0.5 mg/mL) and lactoferrin (0.1-0.3 mg/mL). Based on the generated whey glycoprofile, GlyCAM-1 was the highest contributor of (sialylated) glycans. GlyCAM-1 is not only present in the milk of typical dairy livestock (i.e. bovine, ovine and caprine (Sørensen *et al.*, 1997), but also has been reported in murine milk (Dowbenko *et al.*, 1993). While a GlyCAM-1 gene homologue is found in humans, no functional GlyCAM-1 proteins are secreted into the milk (Rasmussen *et al.*, 2002). In contrast, Gustafsson *et al.* tentatively reported GlyCAM-1 (lactophorin) in human milk, based on the SDS-PAGE derived molecular weight analysis of major glycoproteins (Gustafsson *et al.*, 2005). Trace amounts of GlyCAM-1 were detected by Hettinga *et al.*, also suggesting that a homologue in fact is present in human milk (Hettinga *et al.*, 2011).

Regarding the origin and function of GlyCAM-1, there is limited and also contradicting information. Originally, GlyCAM-1 was described as PP3, a protein originated from the milk fat globule membrane (MFGM). This conclusion was supported by the cross-reactivity of GlyCAM-1 with an antibody to soluble glycoprotein (SGP), an MFGM protein (Kanno, 1989). The conclusion that PP3 is a MFGM protein was questioned in detail by Girardet *et al.*, who instead suggested that the cross reactivity of the anti-SGP antibodies detected GalNAc, an epitope that is common on MFGM proteins, as well as on GlyCAM-1 (Girardet & Linden, 1996). Here, we report that GalNAc, in LacdiNAc motifs, is very abundant on most whey glycoproteins, which may explain the observed cross-reactivity. Contrarily, Bak *et al.* described that a C-terminal peptide of PP3 acted as a membrane anchor (Bak *et al.*, 2000), supporting membrane association. Another publication suggested that GlyCAM-1 (PP3) exists in a membrane bound and a secreted form in mice (Dowbenko *et al.*, 1993). This also fits the observations of Sørenson *et al.* and Hettinga *et al.,* showing significant levels of GlyCAM-1 in both serum, as well as the MFGM fraction (Hettinga *et al.*, 2011; Sørensen *et al.*, 1997).

GlyCAM-1 is not solely expressed in the mammary glands, it is also detected in the epithelial cells of lymph nodes, lungs, uterus and cochlea (Hou *et al.*, 2000). In lymph nodes, GlyCAM-1 mediates lymphocyte trafficking, while in other tissues the function remains unknown. In the mammary gland, GlyCAM-1 expression is regulated by progesterone and prolactin (Hou *et al.*, 2000). Suggested functions for GlyCAM-1 included the inhibition of lipases (Cartier *et al.*, 1990), acting as a lubricant (Dowbenko *et al.*, 1993), or protection against mastitis (Groenen *et al.*, 1995) possibly by its anti-bacterial properties (Campagna *et al.*, 2004). In the immune system, GlyCAM-1 has been indicated in L-selectin mediated leukocyte rolling and trafficking (Springer, 1995). Changes in GlyCAM-1 expression levels have been implicated in inflammation response (Boehmer *et al.*, 2010; Smolenski *et al.*, 2014).

Our data shows that GlyCAM-1 is the major contributor to the mature bovine whey *N*-linked glycoprofile. This novel insight comes as a surprise, as GlyCAM-1 has remained a significantly under-studied protein. Literature concerning this particular protein is relatively scarce and is further complicated due to the different synonyms used for GlyCAM-1, such as lactophorin and PP3. Furthermore, information is contradictory with regards to GlyCAM-1 nomenclature, size, location in the milk, concentration and functionality. GlyCAM-1 is the dominant protein in heated milk; therefore, it is also likely to remain intact in processed whey powders. The proteins lactoferrin and IgG have both been identified as proteins with important immune stimulatory functions, which are mediated by their glycans. The highly glycosylated GlyCAM-1 potentially also has significant effects on the functional properties of the products it is processed into. Lactoferrin *N*-glycans were previously shown to influence TLR-mediated response in THP-1 and HEK293 reporter cell lines (Figueroa-Lozano *et al.*, 2018). Considering the high levels of GlyCAM-1 glycans in bovine milk, this protein is an interesting target for further studies. In conclusion, the approaches reported in this paper for bovine whey glycoprofile analysis allow a rapid screening and interpretation of milk and whey (product) samples from various sources, visualizing variations in individual whey protein concentrations based on their signature glycans. In this work we explored (methods for) the qualitative analysis of the overall bovine milk glycoprofiles. The quantitative potential of this approach remains to be explored.

Acknowledgements

We thank Family Arie and Brechtje van Wijk for the colostrum/milk sample collection at Rietveldhoeve Farm. The authors thank Hjalmar Permentier and Walid Maho at the Mass Spectrometry Facility ERIBA (University of Groningen) for their assistance with the mass spectrometry analysis. This work was financially supported by FrieslandCampina, the University of Groningen/Campus Fryslân, and the University of Groningen.

Supplemental material

Exoglycosidase assay procedure

Glycan digestions with exoglycosidase enzymes were performed in 50 mM sodium acetate buffer at pH 5.5 overnight. The following enzymes were used; jack bean α-mannosidase (75 U/mL in 3.0 M$(NH_4)_2SO_4$, 0.1 mM zinc acetate, pH 7.5), Sigma, green coffee bean α-galactosidase (25 U/mL 100 mM sodium phosphate pH 6.5, containing 0.25 mg/ml bovine serum albumin, ProZyme), bovine testis β-galactosidase (5U/mL in 20 mM sodium citrate phosphate, 150 mM NaCl, pH 4.0, ProZyme), *Streptococcus pneumoniae* sialidase, stong preference for α(2-3) linkages (4 U/mL 20 mM Tris-HCl, 25 mM NaCl, pH 7.5, Prozyme), *Arthrobacter ureafaciens* α-sialidase (5U/mL in 20 mM Tris HCl pH 7.5, containing 25 mM NaCl, ProZyme). *Streptococcus pneumoniae* β-*N*-acetylhexosaminidase (no activity towards GalNAc) (40U/mL in 20 mM Tris-HCl, 50 mM NaCl pH 7.5, ProZyme). Jack bean β-*N*-acetylhexosaminidase (activity towards GlcNAc and GalNAc) (50U/mL in 20mM sodium citrate phosphate pH 6.0, Prozyme). Bovine kidney α-fucosidase 2U/mL in 20 mM sodium citrate phosphate, 0.25 mg/ml BSA pH 6.0, Prozyme). After digestion the enzymes were removed by 10 kDa cut-off centrifugal filters (Millipore). The 2-AA labeled dextran calibration ladder was from Waters Chromatography BV (Etten-Leur, the Netherlands).

Table S1. Sample parameters and analysis data of the colostrum and mature milk samples.[a]

	Cow	1			2			3			4		
	Breed (%)[b]	FV-?-HF (62.5-25-12.5)			FV-HF-MRY (50-37.5-12.5)			HF-MRY (87.5-12.5)			SRB-HF-MRY (50-37.5-12.5)		
	Parity	1			2			7			2		
	Lactation period / Sampling time[c]	Milk (L)	[LF] (mg/mL)	LF (g)	Milk (L)	[LF] (mg/mL)	LF (g)	Milk (L)	[LF] (mg/mL)	LF (g)	Milk (L)	[LF] (mg/mL)	LF (g)
Colostrum	0-4 h	N/A	N/A	N/A	3.2	8.70	27.84	5.0	4.01	20.07	2.2	6.62	14.56
Colostrum	8-14 h	3.5	0.09	0.32	2.0	3.39	6.78	6.0	1.01	6.06	3.0	2.97	8.92
Colostrum	~24 h	1.1	0.11	0.12	3.0	1.80	5.40	7.2	0.52	3.72	3.5	1.97	6.89
Colostrum	~36 h	6.2	0.03	0.20	4.0	1.61	6.45	8.0	0.26	2.04	N.A.	N.A.	N.A.
Colostrum	~48 h	6.5	0.03	0.21	5.0	1.35	6.73	9.2	0.20	1.82	4.5	0.98	4.39
Colostrum	~60 h	6.2	0.04	0.24	4.8	0.91	4.36	8.0	0.19	1.52	4.5	0.67	3.02
Colostrum	~72 h	6.5	0.07	0.45	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	4.0	0.56	2.26
Milk	30 d	11.0	0.03	0.28	7.5	0.19	1.41	12.0	0.16	1.97	7.5	0.11	0.80
Milk	60 d	11.0	0.02	0.26	8.0	0.21	1.68	16.5	0.10	1.62	7.6	0.08	0.63
Milk	90 d	11.0	0.06	0.67	9.5	0.22	2.04	15.0	0.19	2.90	8.5	0.11	0.95
	Cow	5			6			7			8		
	Breed (%)[b]	HF-MRY (50-50)			HF (100)			?			HF (100)		
	Parity	7			3			2			2		
	Lactation period / Sampling time[c]	Milk (L)	[LF] (mg/mL)	LF (g)	Milk (L)	[LF] (mg/mL)	LF (g)	Milk (L)	[LF] (mg/mL)	LF (g)	Milk (L)	[LF] (mg/mL)	LF (g)
Colostrum	0-4 h	3.7	1.56	5.76	4.5	1.69	7.59	4.5	2.68	12.06	10.0	0.48	4.76
Colostrum	8-14 h	4.2	0.51	2.14	4.0	0.46	1.84	4.0	0.86	3.42	6.5	0.15	0.99
Colostrum	~24 h	7.0	0.31	2.18	7.0	0.24	1.71	4.0	0.37	1.46	7.0	0.11	0.76
Colostrum	~36 h	7.5	0.21	1.55	9.0	0.15	1.34	6.0	0.21	1.23	7.5	0.10	0.74
Colostrum	~48 h	9.0	0.15	1.35	9.5	0.11	1.01	6.5	0.14	0.89	9.5	0.08	0.78
Colostrum	~60 h	11.0	0.14	1.53	N.A.	N.A.	N.A.	7.0	0.10	0.72	11.0	0.06	0.66
Colostrum	~72 h	N.A.	N.A.	N.A.	10.5	0.08	0.80	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Milk	30 d	13.0	0.06	0.79	15.0	0.07	1.05	10.0	0.03	0.30	15.5	0.04	0.62
Milk	60 d	12.5	0.07	0.88	16.4	0.05	0.84	10.0	0.05	0.54	14.0	0.08	1.12
Milk	90 d	15.0	0.10	1.49	15.5	0.09	1.44	9.0	0.03	0.29	14.2	0.08	1.12

[a]Colostrum (Col), and milk samples were collected from 8 cows of the Rietveldhoeve Farm, Aduard, Groningen.
[b]Abbreviations used for the genetic background of the cows: Swedish red (SRB), Holstein Friesian (HF), Meuse-Rhine-Yssel (MRY), Fleckvieh (FV) and unknown (?).
[c]Sampling times, lactoferrin concentrations, total amounts of lactoferrin per milking and parity at the time of the first lactation are shown. Time is specified in hours (h) or days (d), respectively. Milk sample volumes are shown in liters (L). N/A: sample not available. Table adapted from Valk-Weeber *et al.* (2020a).

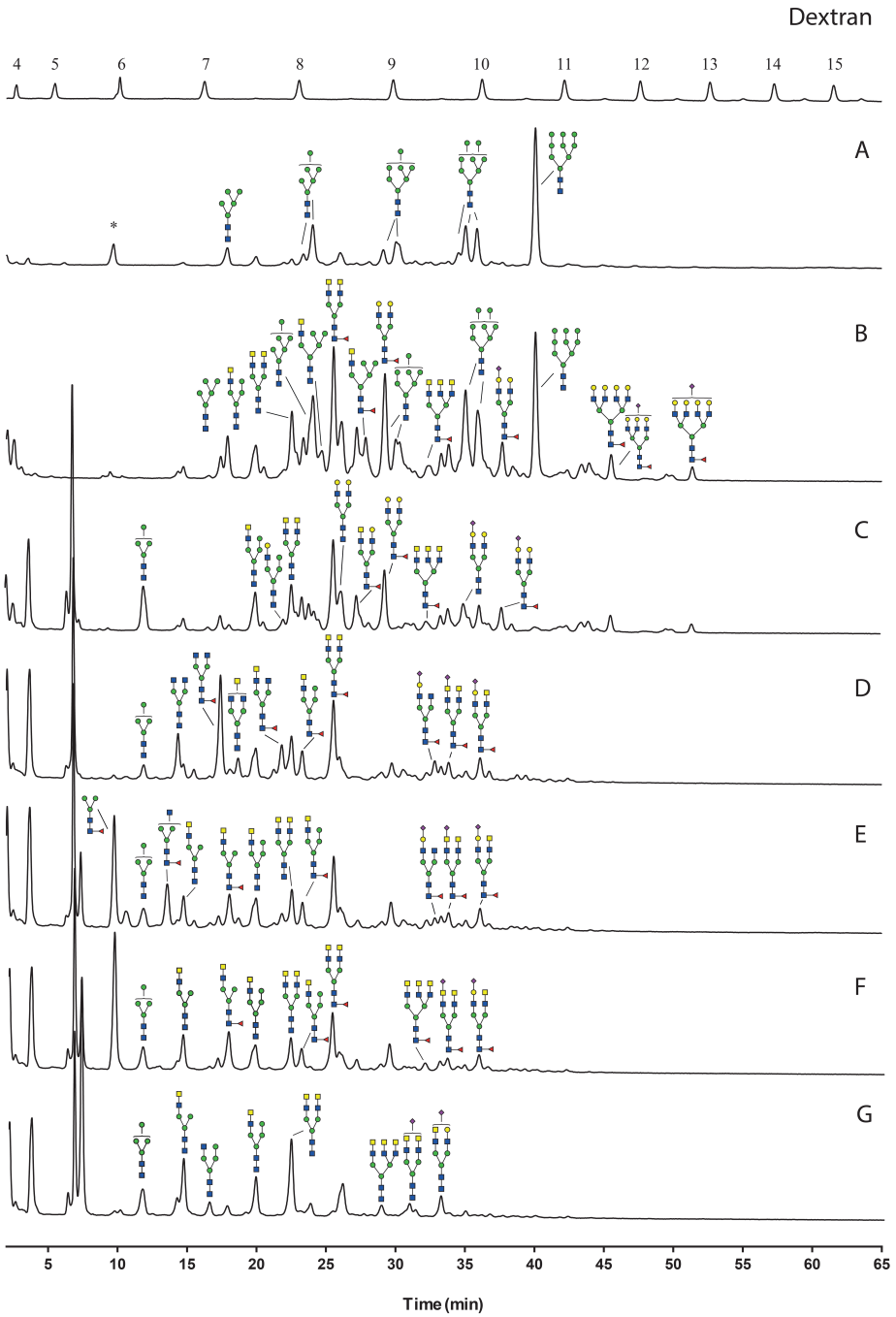


Figure S1. Exoglycosidase assay of the neutral (2-AA labeled) glycan fraction of a bovine whey glycoprofile. For reference, Dextran (top), Lactoferrin (A) and undigested neutral glycan fraction (with a small amount of carry-over from sialylated glycans) (B). Sequential digestion with jack bean α-mannosidase (C), bovine testis β-galactosidase (D) *Streptococcus pneumoniae* β-*N*-acetylhexosaminidase (no activity towards GalNAc) (E) jack bean β-*N*-Acetylhexosaminidase (activity towards GlcNAc and GalNAc; note, incomplete digestion) (F) Bovine kidney α-fucosidase (G).

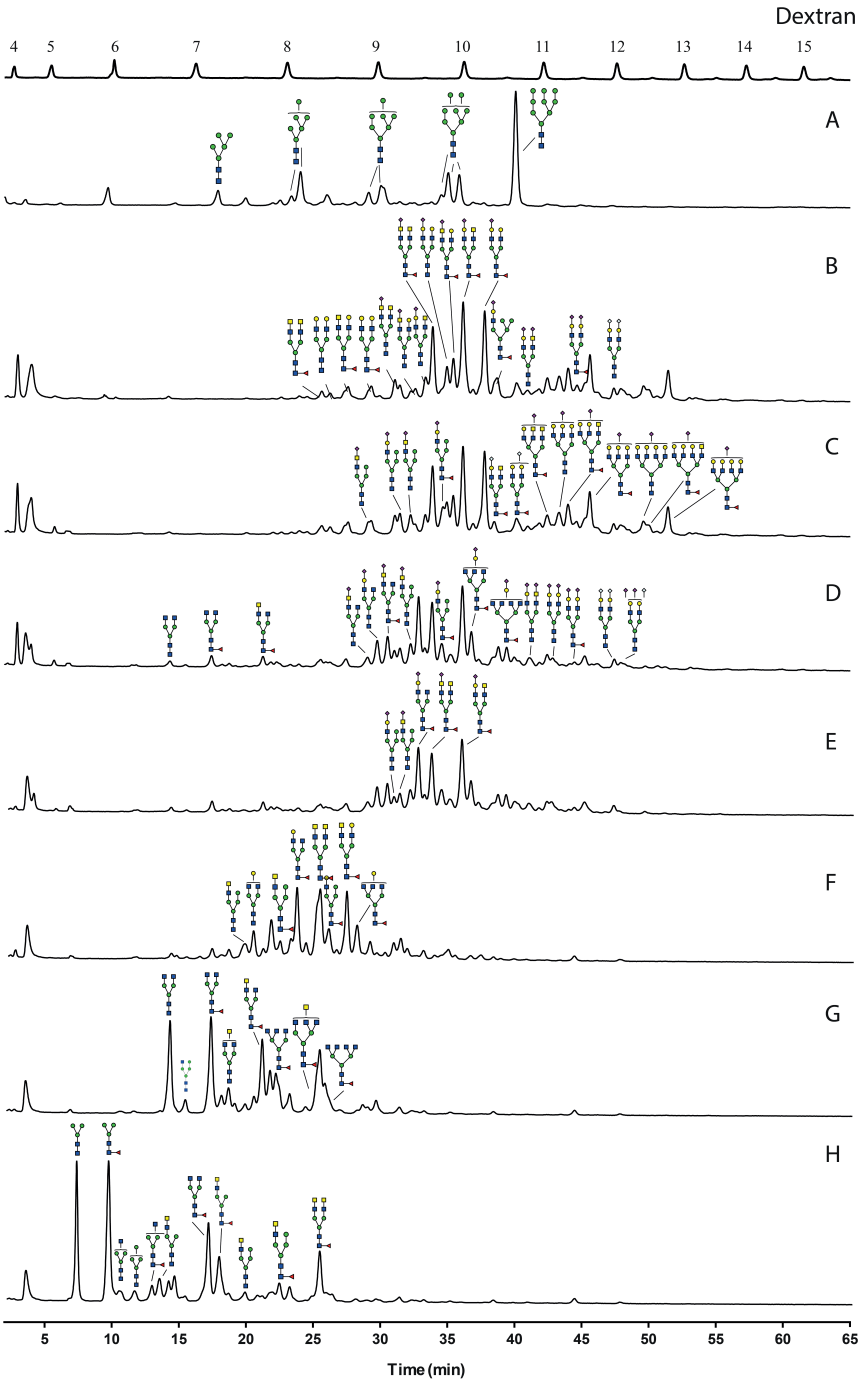
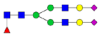
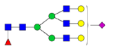
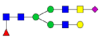
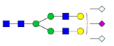
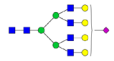
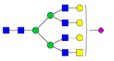
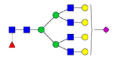
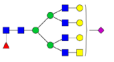
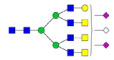
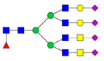


Figure S2. Exoglycosidase assay of the acidic (2-AA labeled) glycan fraction of a bovine whey glycoprofile. For reference, Dextran (top), Lactoferrin (A) and undigested acidic glycan fraction, (with a small amount of carry-over from neutral glycans) (B). Sequential digestion with jack bean α-mannosidase (C), bovine testis β-galactosidase (D) *Streptococcus pneumoniae* sialidase (E) *Arthrobacter ureafaciens* α-sialidase (F) bovine testis β-galactosidase (G) *Streptococcus pneumoniae* β-*N*-acetylhexosaminidase (no activity towards GalNAc) (H).

Table S2. Overview of the bovine whey glycoprotein *N*-glycan structures detected by mass spectrometry analysis. Glycans were labeled with 2-AB prior to analysis. Structures are marked with * when detected as $[M + 2H]^{2+}$.

No.	Theoretical m/z (+2-AB) $[M + H]^{+}$ or $[M + 2H]^{2+}$(*)	Found m/z (+2-AB) $[M + H]^{+}$ $[M + 2H]^{2+}$	Structure	No.	Theoretical m/z (+2-AB) $[M + H]^{+}$ or $[M + 2H]^{2+}$(*)	Found m/z (+2-AB) $[M + H]^{+}$ $[M + 2H]^{2+}$	Structure
1	1355.5095	1355.5080		**19**	1843.7214	1843.7175	
2	1437.5626	1437.5581		**20**	1849.6843	1849.6775	
3	1437.5626	1437.5540		**21**	1890.7109	1890.7058	
4	1517.5624	1517.5582		**22**	1907.7262	1907.7141	
5	1558.5889	1558.5857		**23**	1907.7262	1907.7147	
6	1583.6205	1583.6139		**24**	1923.7211	1923.7102	
7	1599.6154	1599.6110		**25**	1948.7527	1948.7575	
8	1599.6154	1599.6040		**26**	1989.7793	1989.7676	
9	1640.6419	1640.6318		**27**	2003.7208	2003.7167	
10	1679.6152	1679.6112		**28**	1006.3722*	1006.3633*	
11	1720.6417	1720.6370		**29**	1018.8881*	1018.8880*	
12	1745.6734	1745.6643		**30**	1026.8855*	1026.8862*	
13	1745.6734	1745.6628		**31**	1026.8855*	1026.8853*	
14	1761.6683	1761.6625		**32**	1034.8830*	1034.8803*	
15	1761.6683	1761.6625		**33**	1047.3988*	1047.3991*	
16	1786.6999	1786.6684		**34**	1047.3988*	1047.4003*	
17	1802.6948	1802.6841		**35**	1055.9065*	1055.8994*	
18	1841.6680	1841.6627		**36**	1063.9083*	1063.8987*	

No.	Theoretical m/z (+2-AB) $[M + H]^+$ or $[M + 2H]^{2+}$(*)	Found m/z (+2-AB) $[M + H]^+$ $[M + 2H]^{2+}$	Structure	No.	Theoretical m/z (+2-AB) $[M + H]^+$ or $[M + 2H]^{2+}$(*)	Found m/z (+2-AB) $[M + H]^+$ $[M + 2H]^{2+}$	Structure
37	1067.9121*	1067.9070*		**55**	1245.4622*	1245.4549*	
38	1079.4012*	1079.4029*		**56**	1250.4782*	1250.4909*	
39	1099.9144*	1099.9109*		**57**	1261.4571*	1261.4534*	
40	1099.9144*	1099.9121*		**58**	1265.9754*	1265.9722*	
41	1107.9119*	1107.9074*		**59**	1282.4806*	1282.4772*	
42	1107.9119*	1107.9074*		**60**	1302.9938*	1302.9770*	
43	1120.4277*	1120.4146*		**61**	1319.4989*	1319.4862*	
44	1120.4277*	1120.4267*		**62**	1323.5071*	1323.5054*	
45	1136.9329*	1136.9463*		**63**	1325.9784*	1325.9749*	
46	1140.9410*	1140.9405*		**64**	1333.9759*	1333.9685*	
47	1157.4461*	1157.4390*		**65**	1341.9733*	1341.9623	
48	1172.4332*	1172.4326 *		**66**	1392.0177*	1392.0095*	
49	1177.9594*	1177.9478*		**67**	1412.5309*	1412.5145*	
50	1180.4307*	1180.4284*		**68**	1465.0467*	1465.0405*	
51	1188.4281*	1188.4273*		**69**	1485.5599*	1485.5507*	
52	1198.4726*	1198.4691*		**70**	1752.6504*	1752.5848*	
53	1209.4516*	1209.4457*		**71**	1942.7163*	1942.636*	
54	1229.9649*	1229.9610*					

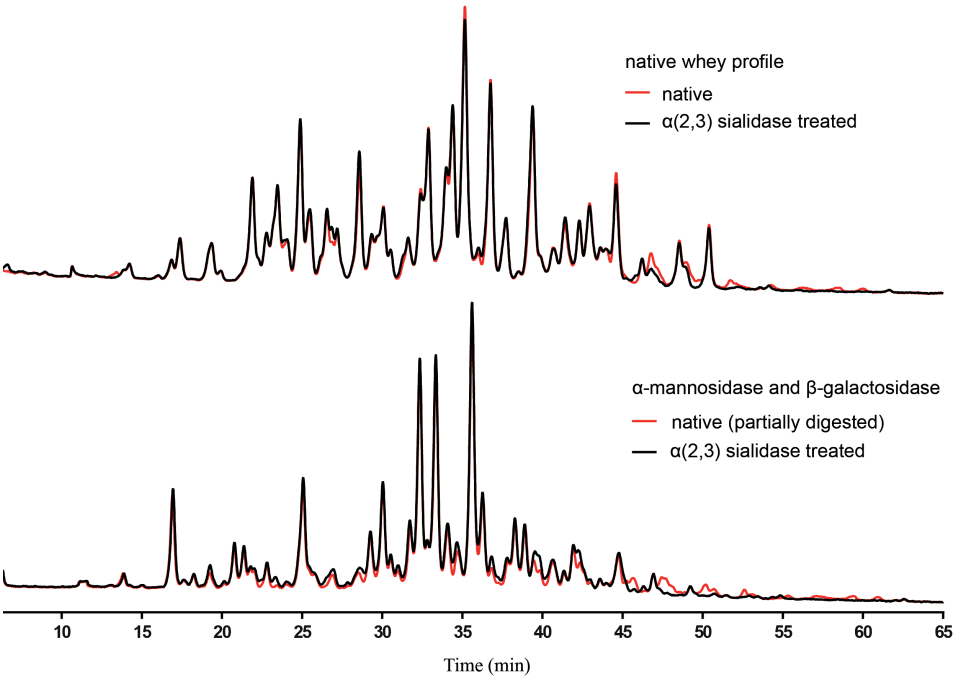


Figure S3. Whey glycoprofiles before and after digestion with Streptococcus pneumoniae sialidase α(2,3)-specific. Top: digestion of an untreated whey sample. Bottom after prior digestion with jack bean α-mannosidase and bovine testis β-galactosidase.

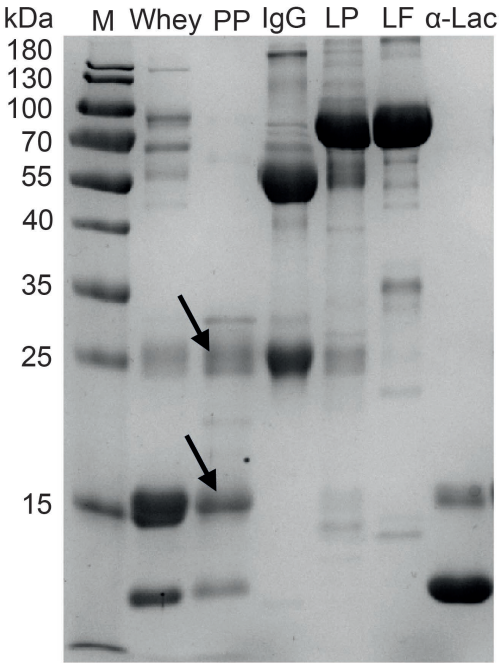


Figure S4.
SDS-PAGE analysis of whey and glycoproteins on a 12.5% polyacrylamide gel. Lanes are marked as marker (M), Acid whey (Whey), proteose peptone fraction (PP, main protein GlyCAM-1, doublet marked with arrows), Immunoglobulin G (IgG), Lactoperoxidase (LP) and Lactoferrin (LF) and α-lactalbumin (α-Lac). ~5 µg total protein on gel for all lanes except PP, which was added from undiluted heated whey (~1 µg PP on gel). In acid whey, the most prominent bands were from β-lactoglobulin and α-lactalbumin, with minor bands from lactoferrin, IgG and GlyCAM-1. The SDS-PAGE analysis of heated whey PP showed the doublet of GlyCAM-1, running at approximately 28/18 kDa. In addition, PP5 (14.3 kDa) was observed 37. Lactoferrin (78 kda), IgG (57/26 kDa) and α-lactalbumin (non-glycosylated 14.1 kDa, glycosylated 16 kDa) were sufficiently removed from the PP sample.

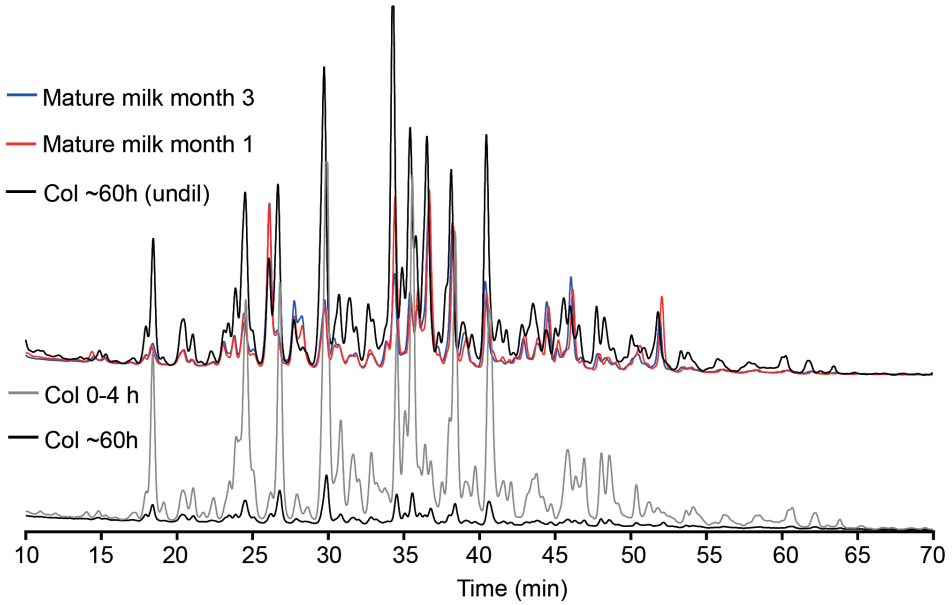


Figure S5. Whey glycoprofiles of Cow 2, 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top).

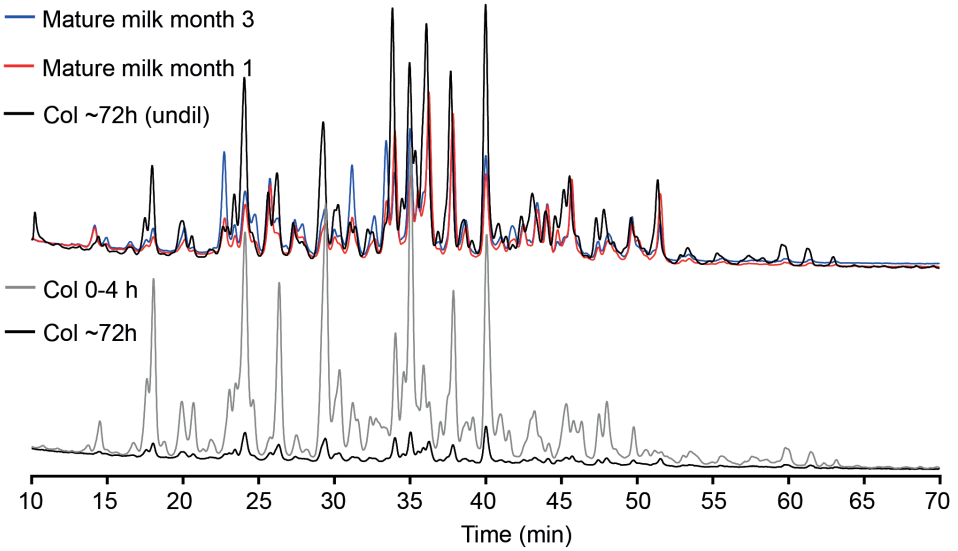


Figure S6. Whey glycoprofiles of Cow 4, 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top).

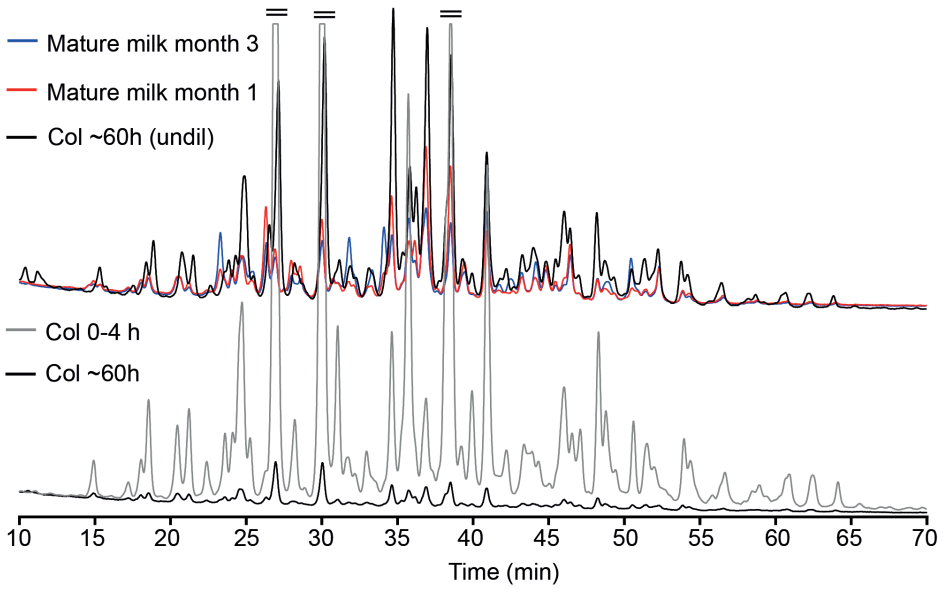


Figure S7. Whey glycoprofiles of Cow 5, 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top).

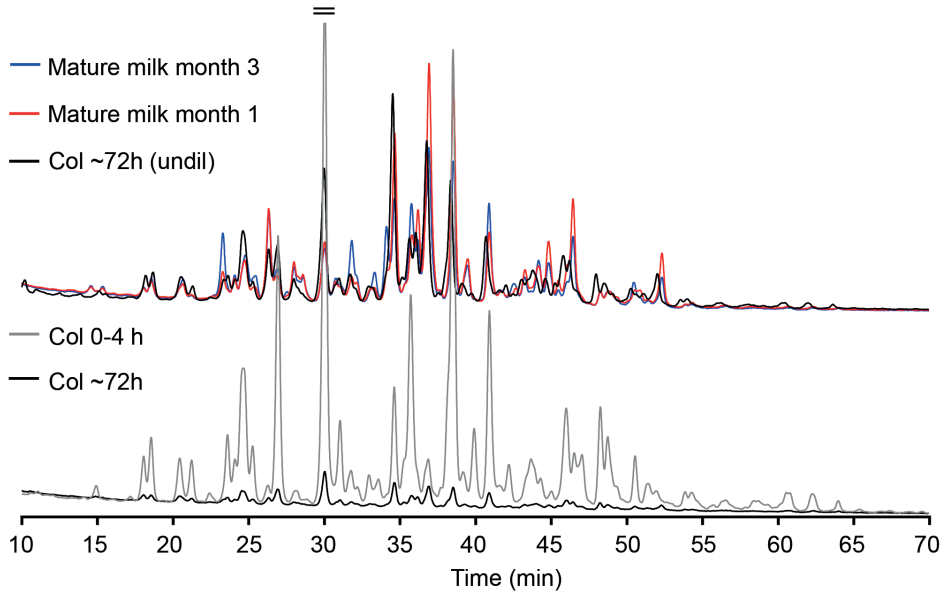


Figure S8. Whey glycoprofiles of Cow 6, 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top).

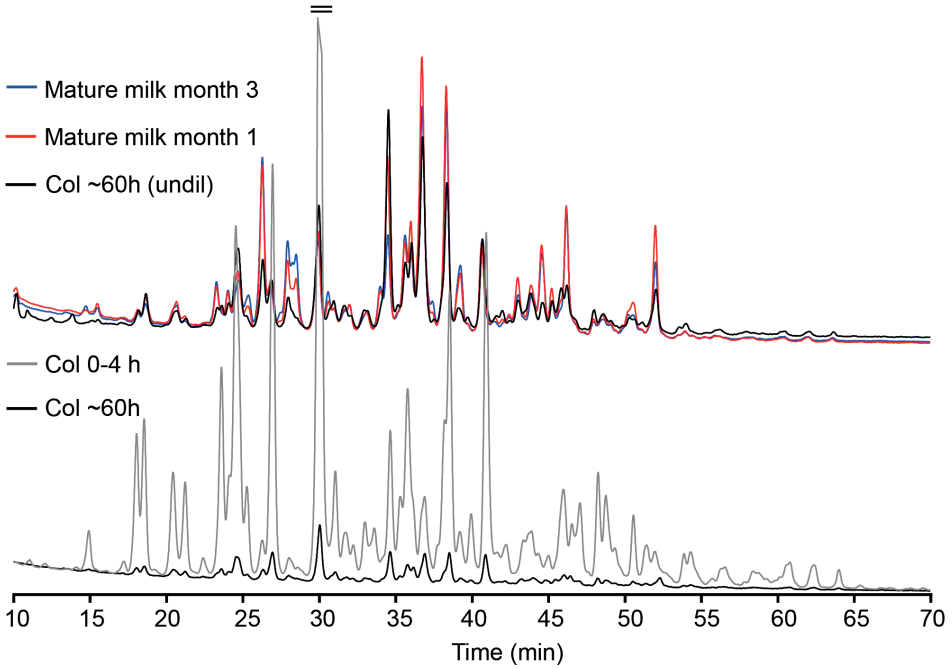


Figure S9. Whey glycoprofiles of Cow 7, 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top).

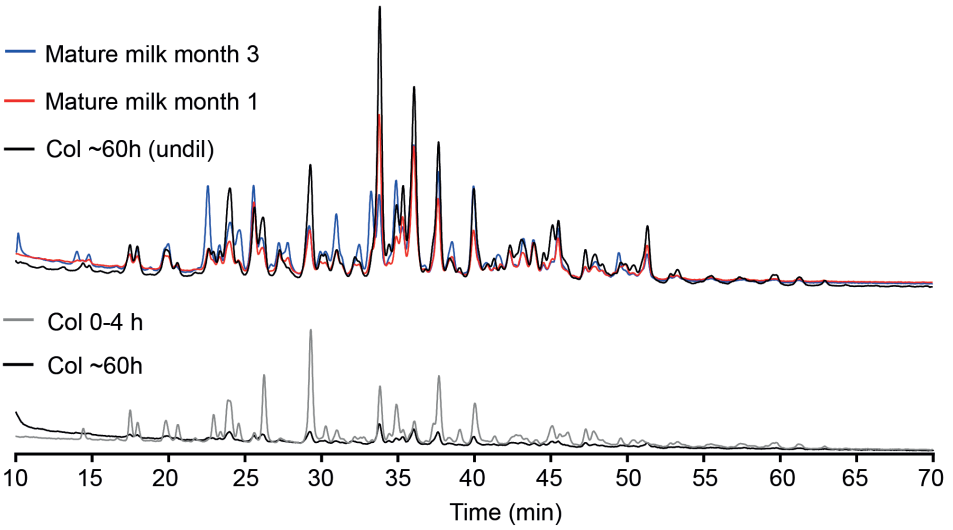


Figure S10. Whey glycoprofiles of Cow 8, 8-times diluted colostrum (bottom) and undiluted (undil.) colostrum versus mature milk (top).